An Evaluation of New After-Action Review Tools in Exercise Black Skies 10 & Exercise Black Skies 12

Julian Vince and Christopher Best

Aerospace Division
Defence Science and Technology Organisation

DSTO-TN-1216

ABSTRACT

An evalution of a new suite of after action review tools was undertaken as part of a series of synthetic collective training research exercises (Exercise Black Skies 10 & Exercise Black Skies 12). This technical note provides a description of the observed benefits of this tool suite and proposes opportunities for future development.

RELEASE LIMITATION

Approved for public release

APPROVED FOR PUBLIC RELEASE

An Evaluation of New After-Action Review Tools in Exercise Black Skies 10 & Exercise Black Skies 12

Executive Summary

Exercise Black Skies 10 (EBS10) and Exercise Black Skies 12 (EBS12) were the latest in a series of simulated collective training exercises conducted by the Defence Science and Technology Organisation (DSTO0 under DSTO task AIR07/327 (Support to Air Force Training Capability and Projects), to address client requirements set forth by Air Force Headquarters. These requirements relate to the development of tools and processes for conducting collective synthetic mission rehearsal.

The broad research aim of both EBS10 and EBS12 was to develop and evaluate a simulated Pitch Black training environment along with supporting training technologies, in order to examine the benefits for a collective of teams resulting from virtual preparation for live training exercises. Among the training technologies being evaluated was the Air Warfare Assessment and Review (AWAR) toolset. The AWAR toolset is a test-bench for evaluating technical and procedural aids in the provision of assessment of training audience performance and subsequent feedback through an After-Action Review (AAR) following the training mission.

This report describes an evaluation of the AWAR suite of tools that was undertaken as part of EBS10 and EBS12. The objectives of this evaluation were to: (1) examine what, if any, differences arose between AARs using AWAR or more standard tools; and (2) identify how AWAR tools and processes may be improved for future synthetic mission training and rehearsal.

The report is divided into two studies, coinciding with the evaluation of the EBS10 and EBS12 versions of AWAR respectively.

In Study 1, the EBS10 version of the AWAR tool suite was observed to encourage a more inclusive and discursive AAR, which took account of the factors impacting on participant learning. AWAR also enabled an objective ground truth to be readily available to learners, to overcome the shortcomings of memory of historical events in a training mission. AWAR also appeared to enhance the opportunity for less experienced participants to learn through the observation of more experienced peers.

In Study 2, the EBS12 version of the AWAR tool suite was observed to support a skilful instructor in drawing out issues for discussion and testing training audience knowledge. However, a potential shortcoming of the tool suite may lie in its lack of application to support practice exercises prior to the next training mission.

Suggestions for future development of the AWAR tool suite include developing mechanisms for allowing training audience members to exhibit and practice the application of the knowledge gained through an AAR prior to the next training mission.

Authors

Julian Vince
Aerospace Division

Julian Vince joined DSTO after completing a Master of Industrial/Organisational Psychology at Deakin University. His research work includes exploring methods and issues surrounding training for distributed teamwork.

_____   _____

Christopher Best
Aerospace Division

Christopher Best is a senior research scientist within Aerospace Division's Crew Environments and Training branch. He holds a B.A. (Hons) in Psychology (awarded in 1997) and a PhD in Psychology (awarded in 2001). His research and professional interests include human-machine interaction, team performance, and team training. Dr. Best is Science Team Leader for the collective training component of DSTO task AIR 07/327 – Support to Air Force Training Capability and Projects.

_____   _____

This page is intentionally blank

Contents

1. Introduction

1.1 Background

The use of synthetic training environments for training collectives of teams may hold potential for overcoming some of the shortcomings of large and expensive live training exercises. However, the implementation of simulation-based training environments potentially invokes some complexities unique to these environments. In order to address the potential training risks of large and complex simulation training events, these events and associated toolsets must be trialled and evaluated.

Exercise Black Skies 10 (EBS10) and Exercise Black Skies 12 (EBS12) were the latest in a series of simulated collective training exercises conducted by DSTO under DSTO task AIR07/327 (Support to Air Force Training Capability and Projects), to address client requirements set forth by Air Force Headquarters. These requirements relate to the development of tools and processes for conducting collective synthetic mission rehearsal. The objectives of EBS10 and EBS12 built upon the outcomes of previous exercises such as the Pacific Link exercises (e.g., Best, Hasenbosch, Skinner, Crane, Burchat, Finch, Gehr, Kam, Shanahan & Zamba, 2007), and Exercise Black Skies 08 (e.g., Shanahan, Best, Finch, Stott, Tracey, & Hasenboch, 2009).

Exercise Black Skies 10 (EBS10) attempted to extend the application of synthetic collective training tools and processes to new air warfare contexts. EBS10 comprised two separate human-in-the-loop simulation activities, run in the Air Operations Simulation Centre, DSTO Melbourne. The first week was aimed at the evaluation of simulation-based mission preparation for Close Air Support (CAS) teams, comprising of F/A-18 pilots and Joint Terminal Attack Controllers (JTACs) (Stevens, Crone, Temby, Best & Simpkin, 2011). The second week was aimed at the evaluation of simulation-based mission preparation for an Air Battle Management (ABM) team (Stott, Best, & Shanahan, 2011).

EBS12 attempted to further expand the bounds of Exercise Black Skies simulated collective training events in terms of scale. The expanded scale of the collective simulation training exercise was achieved through the joint exposure of a Royal Australian Air Force (RAAF) 41 Wing Air Defence Ground Environment (ADGE) ABM team and a RAAF 42 Wing Wedgetail mission crew to a simulated Exercise Pitch Black training environment.

The broad research aim of both the second week of EBS10 and EBS12 was to develop and evaluate a simulated Pitch Black training environment along with supporting training technologies, in order to examine the benefits for a collective of teams resulting from virtual preparation for live training exercises. Among the training technologies being evaluated was the Air Warfare Assessment and Review (AWAR) toolset. The AWAR toolset is a test-bench for evaluating technical and procedural aids in the provision of assessment of training audience performance and subsequent feedback through an After-Action Review (AAR) following the training mission.

AARs are important because they give the training audience an opportunity to frame their subjective understanding of mission performance with reference to a more objective 'ground truth' understanding of the events. An AAR ground truth that contains more objective information than the participants had available to them during the execution of the mission is important because perceptions and memories of events can be distorted (Goldberg & Meliza, 1993). AWAR was in part developed to aid in representing an objective ground truth to a training audience.

There were two perspectives taken with regard to the evaluation of the AWAR tool. These were: (1) an evaluation of the usability of the assessment and review functions of the EBS12 AWAR tool, and (2) the impact of the AWAR tool on the AAR learning process. This report deals solely with the evaluation of EBS10 and EBS12 versions of the AWAR tool impact on the learning process. The evaluation of EBS12 AWAR tool usability for assessment and review functions is reported in Vince & Parker (2013).

During the second week of EBS10 and the whole of EBS12 the AWAR tool suite was applied in ABM post-simulation after action reviews (AARs). The suite of tools for each of EBS10 and EBS12 are described in the relevant section of this two-part report. The objectives of this evaluation were to: (1) examine what, if any, differences arose between AARs using AWAR or more standard tools; and (2) identify how AWAR tools and processes may be improved for future simulation-based mission preparation.

1.2 AARs and training feedback

Training feedback can be categorized into two categories: intrinsic feedback and extrinsic feedback (Meliza, Golberg & Lampton, 2008). Intrinsic feedback arises from the cues and signals that guide behaviour in the task environment. For example, intrinsic feedback for an ABM might come from changes on their work display, or from other team members relaying their perspectives on changes in the situation during mission execution. On the other hand, extrinsic feedback is provided outside the task environment, often in the form of an AAR.

The aim of an AAR is to provide participants with a relatively objective ground truth. An AAR ground truth that contains more objective information than the participants had available to them during mission execution is important because perceptions and memories of occurrences can be distorted (Goldberg & Meliza, 1993). This effect is commonly known as the "fog of war". The fog of war arises from the limited perceived experience of the participants in relatively complex, large-scale events.

The use of ground truth within an AAR helps participants to verify and elaborate their perceived experience. Verification takes place, for example, when feedback from an assessor indicates whether an action or response in the mission was appropriate, given the tactical situation. Elaboration, on the other hand, refers to the use of ground truth to highlight relevant cues and to guide the learner toward correct or more appropriate responses. Elaboration information in an AAR may address appropriate responses or discuss particular errors made by the participants (Shute, 2008).

1.3 General Approach

The analyses of the impact of the AWAR tool suite on the learning processes of the EBS10 and EBS12 training audiences is reported as two studies. The reporting of the results of the analysis of each event in a sequential format allows for the description of the development of both the AWAR tool and the understandings developed about the provision of feedback to a training audience. The development of the understandings of AAR feedback has necessitated the use of two related levels of analysis.

Study 1 examines the AAR feedback processes from a relatively higher level of abstraction. The study focuses on the pattern of interaction between the assessor leading the AAR and the training audience as mediated by the AWAR toolset. This initial study examines AAR data collected from the second week of EBS10. EBS10 AAR data was analysed through the lens of Activity Theory (AT), which is useful in examining how tools mediate social and cognitive processes (Leont'ev, 1978).

Study 2 examines the AAR feedback process at the level of discussion of a single learning point. This study focuses on the strategies used by the assessor to develop the knowledge of the training audience as mediated by the AWAR toolset. This second study examines AAR data collected from EBS12. EBS12 AAR data was analysed with reference to the Algo-Heuristic Theory (AHT) of learning and instruction (Landa, 1974). AHT was developed on principles consistent with AT, and is a theory of learning and instruction aimed at the development of both expertise and general cognitive skills (Bedny & Meister, 1997).

2. Study 1

2.1 Styles of AAR

The literature on AARs broadly describe two general approaches: the 'traditional debrief' and the Socratic method. A traditional debrief is focused heavily on the observations and evaluations of a Subject Matter Expert (SME) observer who gives their version of the ground truth as they see it (Hoare, 1996). This style may be contrasted with the Socratic approach, in which a series of leading questions are used by a facilitator to help the training audience to discover what happened and why. The traditional debrief and the Socratic method might be thought of as two ends of a continuum of possible approaches to AAR. In reality most debrief events fall somewhere in between these two extremes.

A low level of inclusiveness in AAR discussion can present problems for the learning of the training audience. Problems with the low level of inclusiveness used in the traditional debrief approach are: (1) The SME undertaking the debrief may miss important details present in the experience of the participants. This lack of detail may render the SME's advice incomplete, inaccurate, or unconvincing, and (2) Due to the limited scope for inputs into the debrief, the participants' collective experience may not be effectively used as the ground truth for elaboration information. Therefore, the traditional debrief format

potentially diminishes the opportunity for participants to access and use elaboration information in their learning.

Since it focuses heavily on the evaluations of an SME, the traditional debrief can be characterised as a one-way exchange of information. On the other hand, the Socratic approach is more dialectical in its orientation. The dialectical approach involves a dialogue between two or more people who hold different points of view. The interlocutors attempt to resolve the truth by reasoned argument underpinned by the available evidence. The advantage of the dialectical approach is that rather than relying upon the persuasiveness of one person, persuasion may occur through the input of multiple parties and evidences to the discussion.

The Socratic approach is recognised as an effective teaching method when dealing with adult and adolescent learners (Tredway, 1995; Strong 1996). The approach of the formal Socratic method requires the instructor to apply a strategy called "*elenchus*" to inspire a response in the learner called "*aporia*". *Elenchus* involves questioning the learner's responses to examine whether the beliefs they hold are consistent. In discovering the inconsistency of their beliefs, the learner experiences *aporia*. *Aporia* is a state of both uncertainty and curiosity. In this state the learner realises that they need to develop an improved understanding through seeking new knowledge. The questioning involved in the *elenchus* is aimed at encouraging the learners to actively pursue the best truth available to them (Paraskevas & Wickens, 2003).

AARs that apply an inclusive and discursive approach are more likely to produce elaborative information for participants to apply to their learning, than the traditional debrief (Meliza, et al., 1998). Like the Socratic method, questions from the facilitator about the participants' understanding provide a useful strategy in promoting of the participant evaluation of their beliefs about the training experience. The application of the *elenchus* strategy can be expected to result in more discussion to resolve any *aporia*-type uncertainty. Informational cues from AAR tools may promote the use of these strategies by acting as starting points for questions and discussions, and as verification points for the development of the AAR discussion.

2.2 Co-ordinating sources of feedback

The design of AAR can affect the ability of the training audience to verify and elaborate their understanding of the training experience. The design of the tools that support an AAR may lend themselves more toward a traditional debrief or a Socratic style AAR (Meliza, 1999) depending on the information they provide and the procedural opportunities they afford. For example, a comprehensive playback and scoring system for a training mission is likely to afford more scope for detailed discussion and elaborative information than the few broad points used in a traditional debrief.

In order to evaluate the effectiveness of feedback tools, methods of evaluation should ideally focus on the how those tools mediate the way in which the training audience develops and receives feedback information. Activity theory (Leont'ev, 1978) is ideally placed for this purpose because of its emphasis on examining the way in which tools

mediate learning (Bellamy, 1996). Activity theory provides guidance in the analysis of the mediating relationships of tools and socio-cultural factors involved in a work or learning activity. Activity theory aids analysis by helping identify the theoretically important elements of the context of activity. By understanding the mediating relationships within an activity, an improved understanding of the activity as a whole is possible.

2.3 Description of activity theory

Activity theory has been applied to examining the introduction or modification of educational tools (e.g., Murphy & Rodriguez-Manzanares, 2008; Barab, Barnett, Yamagata-Lynch, Squire & Keating, 2002; Basharina, 2007). However, it has the potential to guide the study of the impact of tools across a broad range of settings. Activity theory provides a broad theoretical framework for describing the structure, development, and context of any social human activity. The unit of analysis in activity theory is the 'activity', which is constituted of a subject (individual or group), one or more objects, artefacts, socio-cultural rules, a wider community and the division of labour. Figure 1 illustrates the model of activity systems as formulated by Engeström (1987).
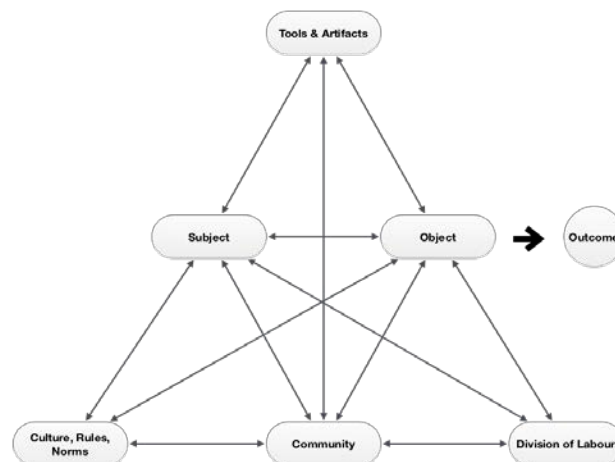
Figure 1. Components of the activity system (adapted from Engeström, 1987)

This framework is useful for analysing the impact of AAR tools because it directs analysis toward how the relationships between the subject (facilitator), community (training audience) and objects (objectives of the AAR) are mediated by the use of physical and cognitive artefacts/tools (e.g. AAR tools and participant knowledge), cultural rules and norms, as well as the division of labour.

There are two important processes drawn from Kaptelinin, Nardi & Macauley's (1999) description of activity theory that aid the analysis of new learning technologies. Firstly, activity theory places great emphasis on the process of mediation. Within the framework, tools are thought to mediate the way in which humans interact with the world through the processes of internalisation and externalisation. Tools may be physical tools for affecting the physical world, or conceptual tools which are used for reasoning. The mediating effects of social rules/norms (culture) and the division of labour are also other mediating elements in activity theory.

Secondly, learning and reasoning involve the transformative processes of tool mediated internalisation and externalisation (Leont'ev, 1978). In the case of internalisation, tools provide a support for learning how to think about a process. For instance a child begins to learn to count by counting fingers or counters, but then internalises this process and loses the need for the tool. Externalisation of internal processes occurs where there is a need to repair or scale a process, such as using pen and paper to keep track of figures in a complex piece of arithmetic. Importantly, activity theorists posit that the form of the tools used in internalisation and externalisation processes shape and constrain the manner in which tasks are performed and problems are solved.

Activity theory therefore provides a way to examine the impact of new AAR tools in terms of what they afford the facilitator and training audience for the development of their learning. Comparing the ways in which the interaction of entities in the activity system is affected by the new and legacy tools highlight how the new tools shape the activity.

2.4 Effects of introducing a new AAR tool suite

According to activity theory, the introduction of a new AAR tool suite should be viewed as an attempt to improve the AAR activity by changing the relations between AAR activity system elements (see Figure 1). The introduction of a new AAR tool suite should change the relationships between the elements of the AAR activity system and these changes should, in turn, result in observable changes in the actions undertaken to achieve the objects of the activity. There are three general kinds of change that could be expected in descending order of granularity: (1) changes in the global mediating effect of the AAR tool on the AAR discussion, (2) changes in the way the objects of the discussion are achieved, (3) changes in the way the AAR tool suite mediates cognitive and social processes (internalisation/externalisation and AAR division of labour). Each of these is described in turn in the sections below.

2.4.1 Global Changes

The first anticipated difference to arise from a new AAR tool suite relates to the mediating effect of the AAR tool suite on the relationship between the facilitator and the training audience. The introduction of a new AAR tool suite should change the interaction between the facilitator (subject) and the training audience (community), relative to the standard AAR. This change may be observable as a difference in the number of contributions generated by the facilitator and the training audience via the AAR tool suite. There are two aspects to this point: (1) an improved AAR tool suite may also encourage a change in the number of contributors to the AAR discussion (inclusiveness); and (2) an improved AAR tool suite may encourage a change in the number of total contributions from the AAR participants (discursiveness).

2.4.2 Object-related Changes

Differences in the mediating relationship between the AAR tool suite, facilitator and training audience may be characterised by examining how the objects of the activity are achieved. AARs conducted during training research events such as EBS10 may be viewed

as having three main objects: (1) improve learners' knowledge, (2) improve technical simulation tools, (3) improve the presentation of white force developed products and entities. The effect of the AAR suite on these objects gives rise to three expected observations.

The first object-related change expected between AARs using the standard and new tool suites relates to how improvement in learner knowledge is achieved. Improvement in learner knowledge is generally achieved through feedback on individual and team mission performance. It is expected that a new AAR tool suite may encourage a change in the inclusiveness and discursiveness of contributions to learner feedback discussions on team and individual mission performance. This means that there will be more contributors making more expansive contributions to the AWAR-based discussion over the standard AAR.

The second object-related change expected between AAR tool suite conditions relates to the impact of technical simulation tool issues on learning. Discussion of the shortcomings of technical simulation tools during AAR represents an attempt to contextualise the learner experience to avoid negative training transfer issues. Such discussion may also indicate where technical simulation tools did not permit the achievement of a desired training experience. It is expected that a new AAR tool suite may encourage a change in the inclusiveness and discursiveness of the contributions to this object.

The third object-related change expected between AAR tool suite conditions relates to issues with white force presentation. White force developed products and controlled entities have an important impact on the fidelity of the mission simulation, and thereby on the development of learner knowledge. Like technical simulation tools, white force presentation issues can detract from learner knowledge development. As a fourth anticipated difference between AAR tool suite conditions, a new AAR tool suite may encourage a change in the inclusiveness and discursiveness of contributions to the discussion of white force presentation issues.

2.4.3 Cognitive and Social Process Changes

There are two other anticipated differences between the AAR tool suite conditions. The first derives from the new AAR tool suite's impact on internalisation/externalisation processes. The second derives from the way actions that support the AAR are divided among the training audience.

Firstly, the two AAR tool suites differ in terms of the way they display mission history information to the training audience. The AAR tool suites are therefore anticipated to differ in terms of the internalisation/externalisation processes they support. Of particular interest is the relative use of participant memory and recorded material of mission history in the development of a ground truth. The processes of internalisation/externalisation should differ between AAR tool suite conditions where there are differing demands on participant memory and access to recorded media. Specifically, introduction of the new AAR tool suite is expected to lead to changes in the inclusiveness and discursiveness of discussion relating to mission history.

Secondly, the way in which labour is divided within the training audience is anticipated to change between the AAR tool suite conditions. The division of labour could be expected to change in two ways: (1) there may be a difference in the amount of labour required to operate the AAR tools, and (2) the AAR tool suites may differ in terms of the way the AAR participants control the AAR tool suite, especially participants of greater relative experience or rank.

2.5 Method

2.5.1 Participants

The participants involved in this study comprised: (1) a Royal Australian Air Force (RAAF) ABM team consisting of five operators, (2) eight RAAF simulation pilot personnel (SIMOPs), (3) two ex-military, white force subject matter experts, (4) a DSTO technical staff member and (5) an ABM team assessor.

The ABM team included an Air Battle Director (ABD), Tactical Director (TD) and three Tactical Controllers (TCs). The SIMOPs supported the ABM team simulation training by providing simulated fast-jet screen entity and voice interaction with the ABM team. The white force subject matter experts, consisting of one former ABM and one former F/A-18 pilot, helped develop and co-ordinate the training scenarios for the ABM team. DSTO-AOD also provided a former ABM member of staff to help run the technical aspects of the AAR. A current 41 Wing ABM posted to DSTO-AOD acted as both a white force SME and as an ABM team assessor.

The TD and assessor were responsible for leading the AARs during the exercise (see Section 2.3.1 below). These participants were comparable in terms of their experience and background in providing AARs. They were contemporaries with experience in the same unit and were accustomed to facilitating AARs using the same procedures. During their careers, the TD and assessor had also been trained and mentored by many of the same instructors.

2.5.2 Materials

2.5.2.1 Video recordings
Digital video/audio recordings were made of six post-mission AARs. A seventh AAR recording was lost due to technical difficulties. The recordings chiefly captured the participants' verbal contributions to the AAR. Inside the visual frame, the video captured the AAR facilitator, large-format screen displays (described in Section 2.2.2 below) and the backs of the ABM team and first few rows of other participants. Many of the SIMOPs sat outside the camera shot, and were therefore difficult to identify personally. However their comments were easily attributable to the SIMOP group.

2.5.2.2 AAR room layout
The recordings were made in an auditorium setting where an AAR facilitator made use of a lectern to speak to several rows of seated participants. To the AAR facilitator's left, there were two large wall-mounted digital screens. These screens displayed mission replay

information and AAR debriefing points. One displayed video capture taken from the ABD/assessor's workstation, or debriefing points from the TD or assessor. The other screen displayed a ground truth picture of the simulated mission. A whiteboard was also used to provide debriefing points and future performance goals during the TD's traditional format AAR.

2.5.2.3 EBS10 Air Warfare Assessment & Review (AWAR) tool suite

The AWAR tool for EBS10 was a prototype training assessment and feedback system that has been evaluated and reconfigured through previous Black Skies exercises. The EBS10 version of the AWAR tool had two components. The first component was a software tool that aimed to provide clear and structured performance assessment and feedback by enabling assessors to score performance and make AAR notes. These notes were time stamped so that they could be linked with video and audio clips of related mission segments. Replays included simulation ground truth, participant screen recordings, and selectable radio and intercom voice recordings. The purpose of this suite of tools was to provide contextually anchored verification of events to reinforce learning points.

Figure 2. AWAR team goal hierarchy, displaying the decomposition of the "Manage Information Systems" goal, sub-goals and related scoring criteria.

The second component of the AWAR tool was a set of team objectives defined by using an approach to task analysis similar to that described by Annett, Cunningham, and Mathias-Jones (2000). These objectives formed a hierarchical decomposition of Defensive Counter Air (DCA) and Offensive Counter Air (OCA) mission goals for the ABM team (Hasenbosch & Best, 2007). These team goal hierarchies were ABM-specific extensions adapted from the Australian Joint Essential Tasks (ASJETs: McCarthy, Kingston, Johns, Gori, Main, & Kruzins, 2003). The DCA and OCA team goals were arranged into a means-ends hierarchy. The lower-level goals and sub-goals of the hierarchy explained *how* the higher level goals and goals categories were to be achieved. The higher-level goals and goal categories explained *why* the lower level goals and goal categories were important to achieve. Criteria associated with each of the lowest-level objectives provided a description of what constitutes good performance on each. The purpose of arranging ABM team goals into such a hierarchy was to structure assessment and feedback in a way that emphasised not only what happened during a mission, but also why it happened, and what the flow-on effects could be.

2.5.3 Procedure

2.5.3.1 AARs

Over the course of the exercise, seven AARs took place. For each of the AARs the participants gathered in the briefing/debriefing space within several minutes after each simulated mission. The initial DCA mission (DCA 1a) was debriefed in a manner consistent with current practice and was therefore used as an example of a standard debrief. This debrief made no use of the AWAR tools suite. The agenda of this "standard AAR" consisted of: (1) an initial introduction by the ABD, (2) an ABM-team mission debrief facilitated by the TD where mission tool deficiencies and ABD/TD-derived objectives were reviewed and evaluated, (3) a simulation tool debrief, and (4) an overview commentary both provided by the ABD. Prior to the ABD/TD-derived objectives review, a faster-than-real-time run-through of mission screen recording was played. This run-through was accompanied by a description of events by the TD.

The second AAR (OCA 1a) was the same as the first, except that the commentary of the mission run-through was undertaken by a visiting SME, as a separate debrief after the main AAR. This semi-standard AAR made no use of the AWAR tool suite.

The remaining AARs made use of the AWAR tools and followed a consistent agenda. The agenda for these "AWAR AARs" was: (1) an initial introduction by the ABD, (2) a traditional review of ABD and TD-derived objectives delivered by the TD, (3) an AWAR debrief facilitated by the assessor including book-marked video and audio replay and performance evaluation scored against the OCA hierarchy, (4) a traditional evaluation of the performance against ABD/TD-derived objectives that was facilitated by the TD, (5) a simulation tool debrief, and (6) an overview commentary, both provided by the ABD.

2.5.4 Data analysis

Audio recordings of the six retrieved AARs were transcribed. The speech events within these transcriptions were then coded via NVivo software (Version 9, 2010; QSR Software,

Doncaster, Victoria). The coding scheme that was developed for this purpose was guided by the Grounded Theory approach (Glaser & Strauss, 1967). Grounded theory is an approach to qualitative research that relies upon the emergence of a fitting theoretical description of the context under examination. The approach in this case entailed reading the AAR transcripts with reference to the original video, while also being guided by the activity, actions and goals decomposition described in activity theory.

The coding schema for speech events was structured according to the objects of the training activity. There were judged to be three main objects: Improving learners, improving simulation tools, and improving white force presentation.

The decomposition of the object "improving learners" is shown in Figure 3. Improving learners was evident in the transcripts via two main actions: (1) discussion of ABM team performance, and (2) discussion of individual performance. The goals of these actions were to: (1) highlight and assess team-related performance issues, and (2) highlight and assess individual performance issues, respectively.

Object of Activity
(Improve Learners)
↓

Actions		*Goals*
(Discuss team performance)	→	Highlight/assess ABM team performance issues
(Discuss individual performance)	→	Highlight/assess individual performance issues
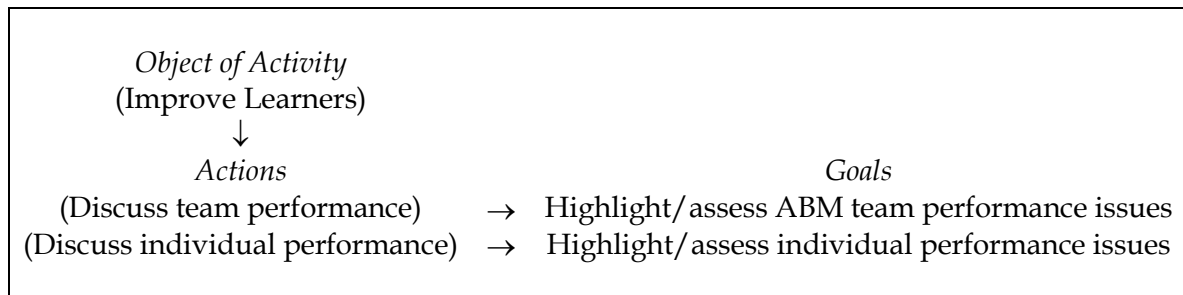
Figure 3. Hierarchical decomposition of the "improve learners" object

The decomposition of the object "improve simulation tools" is shown in Figure 4. Improving simulation tools was evident in the transcripts via one action - discuss technical simulation tool issues. The goals of this action were to review and assess simulation tool impacts on ABM-team learning.

Object of Activity
(Improve Simulation Tools)
↓

Actions		*Goals*
(Discuss technical sim tools)	→	Review/assess simulation tool impacts on learning
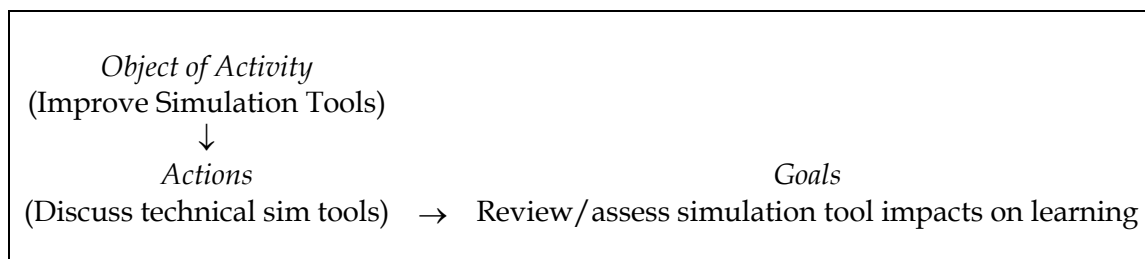
Figure 4. Hierarchical decomposition of the "improve simulation tools" object

The decomposition of the object "improve white force presentation" is shown in Figure 5 Improving white force presentation was evident in the transcripts via two main actions: (1) discussion of military routine presentation, and (2) discussion of simulated air entities (SIMOP) presentation. Military routine presentation included the accuracy of white force derived products and processes like kneepad cards or whether necessary elements were included in the pre-mission brief. The goals of these actions were to: (1) review and assess

military practice or routine issues and their impact on ABM-team learning, and to (2) review/assess SIMOP presentation issues and their impact on ABM-team learning, respectively.

Object of Activity
(Improve White force
Presentation)
↓

Actions		*Goals*
(Discuss military routine presentation)	→	Review/assess military routine presentation impacts on learning
(Discuss SIMOP presentation)	→	Review/assess SIMOP presentation impacts on learning

---
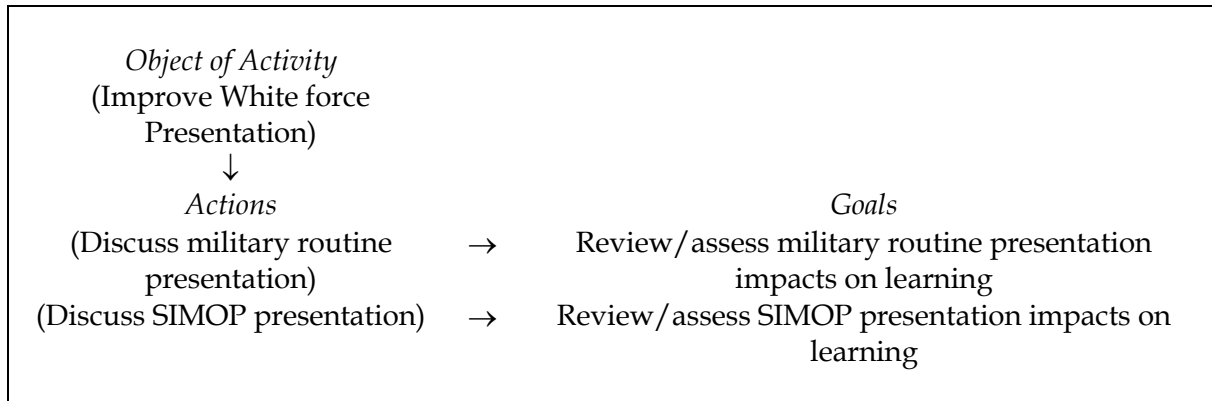
Figure 5. Hierarchical decomposition of the "improve white force presentation" object

A fourth set of categories was also developed at the action level. These categories were developed for coding observed speech events of a more general nature. These categories were: (1) requests for AAR tool manipulation by the DSTO staff, (2) descriptions of mission history that were not evaluative in purpose and, (3) 'meta-discussion' events, that is, speech events that were about the discussion itself. Meta-discussion events were typically speech events that helped orient participant to topic changes within the AAR discussion.

Data analysis focussed on the TD's mission debrief in the DCA 1a[1] and OCA 1a "standard AARs". For the four "AWAR AARs" that followed, the assessor's AWAR debrief was the section that was analysed. The timing of the presentation of the standard AARs before the AWAR AARs was driven by a concerns for order effects. While there was a possibility that the standard AARs may not have been an ideal representation of standard AAR, due to their placement early in the exercise, when participants' roles and processes are bedding down, a greater concern was the potential contamination of the standard AAR if it occurred after the AWAR debriefs. Therefore, the decision was made to run the standard AARs first, followed in later days by the AWAR assisted AARs. Comparison between the standard AAR and AWAR AAR conditions were made in terms of frequency counts of speech events falling into the categories described above, as well as qualitative content analysis.

---

[1] Missions have been labeled according to whether they were Defensive Counter Air (DCA) or Offensive Counter Air (OCA), as well as a combined number and letter category indicating the mission sequence and whether it was the first or second mission of the day, e.g. OCA 1a or OCA 1b.

2.6 Results & Discussion

2.6.1 Overall differences

Data were first analysed according to the overall number of speech events observed during each AAR. These data are presented in Table 1 below. To emphasise both the absolute differences in the number of speech events between AARs and the relative contribution of each of the participants, data are presented in tabular form. As can seen at the bottom of Table 1, the mission total frequency counts are higher in all the AWAR AARs than standard AARs. This means that AWAR AARs produced more distinct discussion points than the standard AARs. The OCA 3b was somewhat shorter than the other AWAR AARs. This was because of time pressure to end the AAR and wrap up the training research exercise.

Table 1: Overall speech event frequency: Standard and AWAR AARs

Overall Coded Speech Events	DCA 1a (Standard AAR)	OCA 1a (Standard AAR)	OCA 2a (AWAR)	OCA 2b (AWAR)	OCA 3a (AWAR)	OCA 3b (AWAR)	Total by Role
ABD	4	10	24	17	37	17	109
Assessor	-	2	98	92	103	61	356
DSTO	-	-	8	1	-	3	12
Simops	2	7	2	23	5	2	41
TC1	5	5	6	3	19	-	38
TC2	-	-	5	-	1	1	7
TC3	-	3	2	3	1	-	9
TD	37	36	8	14	19	3	117
W/Force	-	20	20	8	7	7	62
Mission Total	48	83	173	161	192	94	751

There was also support for the expectation that the AWAR AARs would involve more participants in the discussions than the standard AAR. The frequency counts within Table 1 indicate that while the assessor made a greater overall number of comments in the AWAR AARs than the TD in the standard AARs, there were also more contributors to the discussions. The average number of silent participants during the AWAR AARs was 1, while the standard AARs averaged 3.5 silent participants. The broader audience contribution to the AWAR AAR supports the expectation that the tools would encourage a more Socratic style discussion of the learning points. The targeted replays of the ground truth served as starting points for discussion. In contrast the standard AAR format involved a greater reliance on the TD and ABD's interpretations of the ground truth.

2.6.2 Object 1: Improve learners

Data were analysed according to the number of speech events related to team or individual mission performance. These data are presented in Table 2. The standard and AWAR AARs differed in terms of the number of speech events aimed at improving

learners. Table 2 shows that the frequency of speech events relating to team and individual performance was greater in the AWAR AARs than the standard AARs. Table 2 also shows that the number of contributors to the discussion also tended to be higher in the AWAR AARs, with the exception of the last training mission.

Table 2. "Improve learners" speech event frequency: Standard and AWAR AARs[2]

Team/ Individual Speech events		ABD	Assessor	DSTO	SIMOPS	TC1	TC2	TC3	TD	White Force	Mission Total
DCA 1a	Team	-	-	-	-	-	-	-	5	-	5
(Standard)	Individual	-	-	-	-	-	-	-	-	-	
OCA 1a	Team	2	-	-	3	-	-	-	2	10	32
(Standard)	Individual	2	-	-	-	5	-	-	3	5	
OCA 2a	Team	10	32	-	-	3	2	1	2	9	110
(AWAR)	Individual	9	21	2	-	3	3	1	6	6	
OCA 2b	Team	6	30	-	7	-	-	2	5	-	84
(AWAR)	Individual	2	17	-	2	3	-	1	6	3	
OCA 3a	Team	14	38	-	-	5	1	-	9	4	123
(AWAR)	Individual	8	24	-	-	12	-	1	5	2	
OCA 3b	Team	2	22	-	-	-	-	-	-	2	46
(AWAR)	Individual	8	6	-	-	-	1	-	1	4	
Role Total		63	190	2	12	31	7	6	44	45	400

A breakdown of the mean number of types of "improve learner" speech events is presented in Table 3. The "raise issue" category refers to any AAR participant raising an issue for discussion. The "discuss issue" category refers to speech events that further a topic raised or is related to feedback or assessments provided by the assessor in AWAR AARs, or the TD in standard AARs. The "discuss solution" category refers to attempts to offer a solution to any problems raised in the AAR. The "provide feedback/assessment" category refers to speech events in which feedback, assessment and advice were offered by the assessor or other AAR participants. What can be seen is that there are greater mean speech events across all of the speech event types in the AWAR AAR relative to the standard AAR. In particular, the proportionately greater mean number of "provide feedback/assessment" and "raise issue" speech events point to these speech events being an important difference between the two AAR types. The reasons for this difference between AARs is described below.

Table 3. Mean "Improve Learners" speech event types: Standard and AWAR AARs

	Raise Issue	Discuss Issue	Discuss Solution	Provide Feedback/Assessment
Standard AAR	2.5	9	2	4.5
AWAR AAR	18.5	37.5	7.25	29.75

Qualitative analysis of the content of the discussion suggests that the greater inclusiveness and discursiveness of the "improve learners" oriented discussion. The assessor would often stop the replay to raise an issue or provide a feedback or assessment point.

---

[2] Note that Table 2 is oriented differently to the other tables in this report (i.e., roles are arrayed along the horizontal axis, rather than the vertical axis). This has been done in order to fit the table onto a portrait-oriented page.

Consistent with the Socratic approach to developing knowledge, these issues were often raised as questions intended to explore individual and team level understandings of important mission events, as in the following examples:

Assessor: "In terms of raid matching, were you aware of that process happening at that time, TD?"

Assessor: "Can you see the decon(fliction) issue is beginning to happen right now?"

After an initial exploration of the issue by the relevant participants, associated issues were often raised within the group for further discussion. Participants often felt inclined to offer their own understanding of the event for comparison. The use of the AWAR as a probe for exploring and developing participant understandings seemed to replace the relatively less inclusive and discursive style of the standard AAR format. The standard AAR tended to encourage a more directive approach. The TD tended to only seek input from the floor when he had difficulty in recalling the details of the ground truth context.

2.6.3 Object 2: Improve simulation tools

Data were analysed according to the number of speech events about issues related to simulation tools. These data are presented in Table 4. There was support for an expected difference between AAR conditions regarding discussion of technical simulation tool issues. The reason simulation tool issues were raised as part of the AAR is because they can impact upon the training experience of the training audience. The results on the discussion of issues related to simulation tools were contingent on the presence of simulation tool problems in the mission. However, Table 4 suggests that where an issue was raised, it received deeper consideration by the group in the AWAR AAR than the standard AAR. Qualitative analysis indicated that in the standard AAR (OCA 1a), a simulation issue was raised by the TD and responded to by the assessor who also helped managed the simulation environment. In the following two training research missions, the AWAR AAR discussion of simulation tool issues widened into a discussion of the training impact.

Table 4. "Improve simulation tools" speech event frequency: Standard and AWAR AARs

Simulation Tool Speech Events	DCA 1a (Standard AAR)	OCA 1a (Standard AAR)	OCA 2a (AWAR)	OCA 2b (AWAR)	OCA 3a (AWAR)	OCA 3b (AWAR)	Total by Role
ABD	-	-	1	2	-	-	3
Assessor	-	1	3	1	-	-	5
DSTO	-	-	2	-	-	-	2
Simops	-	-	-	5	-	-	5
TC1	-	-	-	-	-	-	-
TC2	-	-	-	-	-	-	-
TC3	-	-	-	-	-	-	-
TD	-	1	-	-	-	-	1
W/Force	-	-	1	-	-	-	1
Mission Total	-	2	7	8	-	-	17

As an example of this, an issue was raised as to why some calls from the ABM team to the SIMOPs were not receiving the appropriate responses. The issue needed clarification because as part of their role the ABM team need to troubleshoot the unusual behaviour of their own-force entities. A SIMOP clarified:

"…just with the comms, we heard that sound all through the channels. I was trying to turn channels down so I could hear them, but I did miss calls, and it wasn't intentional."

The comms noise was an example of a temporary problem with simulation equipment that impacted on the events in the simulation mission. As can be seen in Table 4, as these issues were rectified or workarounds developed later in the week, they were raised less often.

2.6.4 Object 3: Improve white force presentation

Data were analysed according to the number of speech events about issues related to white force presentation. These data are presented in Table 5. The 'improve white force presentation' object comprised of two components. The first component was aimed at improving the military routine presentation aspects of the exercise. The second component was aimed at improving the SIMOP entity presentation. As can be seen in Table 5, the TD made a greater number of speech events about military routine presentation issues than SIMOP presentation issues in the standard AAR. Conversely, the assessor made a greater number of speech events about SIMOP presentation issues than military routine presentation issues. The SIMOPS also made more speech events in the AWAR AAR on the SIMOP presentation topic. This difference in White Force presentation emphasis (military routine versus Simop presentation) between the AARs was at least in part due to the emphasis on the discussion of occurrences in the replay. The AWAR AAR tended to emphasise SIMOP presentation for the learning purposes of the ABM team.

Table 5. "Improve white force presentation" speech event frequency: Standard and AWAR AARs

White Force Presentation Speech Events	DCA 1a (Standard AAR)		OCA 1a (Standard AAR)		OCA 2a (AWAR)		OCA 2b (AWAR)		OCA 3a (AWAR)		OCA 3b (AWAR)		Role Total
	Military Routine	Simop	Military Routine	Simop	Military Routine	Simop	Military Routine	Simop	Military Routine	Simop	Military Routine	Simop	
ABD	1	1	-	2	-	-	-	1	-	1	-	-	6
Assessor	-	-	1	-	1	-	-	7	-	4	1	3	17
DSTO	-	-	-	-	-	-	-	-	-	-	-	-	-
Simops	1	-	3	1	-	-	-	6	-	3	-	1	15
TC1	1	1	-	-	-	-	-	-	-	1	-	-	3
TC2	-	-	-	-	-	-	-	-	-	-	-	-	-
TC3	-	-	-	3	-	-	-	-	-	-	-	-	3
TD	4	-	7	1	-	-	-	2	-	-	-	-	14
White Force	-	-	2	-	-	-	-	2	-	-	-	1	5
Mission Total	7	2	13	7	1	-	-	18	-	9	1	5	63

Standard AAR discussion of white force presentation issues tended to relate to whether military routine tasks or products were accurate and available to the ABM team. These

issues included things like the correctness of kneepad cards and whether necessary elements were included in the pre-mission brief. The AWAR AAR tended to emphasise the circumstances of the SIMOP presentation for the purpose of informing the ABM learning experience, for example:

White Force SME: "This one didn't go according as White Force had planned. This was supposed to be an ROE problem. We thought he was going to get there untargeted. There was supposed to be a low guy you didn't have all the ROE info on - so how are you going to solve it?"

2.6.5 Speech events related to mission history description

Data were analysed according to the number of speech events related to the description of mission history. These data are presented in Table 6. Speech events related to mission history were descriptions of mission history that were not directly related to discussion learning points. They tended to arise through queries about mission event context. Mission history speech events were anticipated to be lower in the AWAR AAR, because the bookmarked replay was expected to lessen the reliance on establishing a ground truth. Table 6 suggests that there was no clear quantitative support for the anticipated difference between AAR conditions in the frequency of mission history speech events.

Table 6. "Mission history" speech event frequency: Standard and AWAR AARs

Mission History Speech Event	DCA 1a (Standard AAR)	OCA 1a (Standard AAR)	OCA 2a (AWAR)	OCA 2b (AWAR)	OCA 3a (AWAR)	OCA 3b (AWAR)	Total by Role
ABD	2	-	4	5	7	3	21
Assessor	-	-	18	4	11	7	40
DSTO	-	-	-	-	-	-	-
Simops	1	-	2	3	2	1	9
TC1	3	-	-	-	1	-	4
TC2	-	-	-	-	-	-	-
TC3	-	-	-	-	-	-	-
TD	13	5	-	-	2	-	20
W/Force	-	-	2	2	1	-	5
Mission Total	19	5	26	14	24	11	99

A breakdown of the mean mission history speech event categories is shown in Table 7. The confirm/amend category relates to speech events that resulted from a need to confirm or amend a mission history description or query.

Table 7. Mean mission history related speech events: Standard and AWAR AARs.

	Confirm/Amend	Description	Query
Standard AAR	3.5	7	1.5
AWAR AAR	4.75	9.5	5

While each mean mission history speech event was higher for the AWAR AARs relative to the standard AARs, the query category was proportionately larger in the AWAR AAR than the confirm/amend and description categories.

Qualitative analysis suggests that there are two main reasons for this: (1) the assessor used mission history queries to check whether some issues needed to be followed up as learning points, and (2) the AWAR was used to replay selected parts of the mission rather than an entire mission flow. Therefore, mission history queries were sometimes used as a check on where the mission was up to in terms of events, or where the mission events were heading next. For example:

Assessor: "Both arms out of this end up kill removing, confirm? In the east?"

2.6.6 Speech events related to AAR tool manipulation

Data were analysed according to the number of speech events related to the manipulation of the AAR tools. These data are presented in Table 8. Speech events related to AAR tool manipulation were instances where a participant asked for the DSTO technical staff to control an AAR replay or presentation tool on their behalf.

Table 8. "AAR tool manipulation" speech event frequency: Standard and AWAR AARs

AAR Tool Manipulation	DCA 1a (Standard AAR)	OCA 1a (Standard AAR)	OCA 2a (AWAR)	OCA 2b (AWAR)	OCA 3a (AWAR)	OCA 3b (AWAR)	Total by Role
ABD	-	-	-	1	6	4	11
Assessor	-	-	19	14	16	10	59
DSTO	-	-	4	1	-	3	8
Simops	-	-	-	-	-	-	-
TC1	-	-	-	-	-	-	-
TC2	-	-	-	-	-	-	-
TC3	-	-	-	-	-	-	-
TD	4	1	-	1	3	2	11
W/Force	-	-	-	-	-	-	-
	-	-	-	-	-	-	-
Mission Total	4	1	23	17	25	19	89

The frequency of AAR tool manipulation requests was collected to assess: (1) how much labour overhead was required in manipulating the AAR tools, and (2) the degree to which the training audience were willing to engage with the AAR tools by requesting manipulations themselves.

The frequency of assessor-to-DSTO technical staff interaction is shown in Table 8. These numbers point to a control overhead in the use of the AWAR tool suite. This highlights a potential area for development for future versions of the tools. Future versions of the AWAR may benefit from incorporating controls that are easier for the assessor to access and use, resulting in a more seamless presentation.

What can also be seen in Table 8 is that AWAR AARs had a greater number of participants requesting AAR tool manipulations. The increased participant requests for AWAR AAR tool manipulations indicated an increased interaction between the assessor, ABD and TD

and the AWAR tool suite. The AWAR tool manipulation speech events tended to occur where more senior member of the ABM team felt that altering the perspective of the AWAR tool display in order to highlight or add to the available information would increase their understanding of events. For example:

ABD: "Can I get a range assessment from Apache Zulu to Osprey at this time, please?"

ABD: "Can I quickly see the ABD's scope?"

TD: "Request to zoom in on Cobra to the southeast there?"

ABD: I would like to hear the content of what's on there?"

A possible benefit of this increased AAR tool interactivity is that junior participants have an increased opportunity to have their perceived ground truth validated, amended and elaborated through observing more experienced participants interact with the AWAR tool suite.

2.7 Study 1 Conclusions

Study 1 aimed to evaluate how the introduction of a novel AAR tool changed the way in which ABM AARs were conducted, in terms of the pattern of interaction between AAR participants. The study has made use of activity theory for this purpose. Activity theory places emphasis on longitudinal change, especially change mediated by the development of tools. Examining the application of the AWAR AAR tool during EBS10 through the lens of activity theory has enabled a developmental comparison with traditional debriefing technology. Several conclusions can be drawn from this comparison.

Firstly, the AWAR tool suite appears to encourage a more inclusive and discursive AAR. The AWAR tool suite seemed to encourage a more Socratic style of debrief, which posed questions where issues were not clear-cut, and made use of participant inputs to elaborate understanding. This relatively Socratic style of AAR was apparent in all the AWAR AARs. This style of AAR was also able to be applied to all of the identified objects of the training research exercise, that is: (1) improving learners, (2) improving simulation tools, (3) improving white force products/processes and SIMOP presentation.

Secondly, the AWAR tool suite enabled an objective ground truth to be readily available to the AAR participants. The AWAR tool suite enabled the important portions of the mission to be replayed coupled with feedback highlighting the relevant learning points. Access to an accurate ground truth to compare to the participant perceived truth has been suggested as important for learner development (Meliza, Golberg & Lampton, 2008).

Thirdly, there are possible social learning advantages associated with the way in which the AWAR increases the opportunity for interaction with tools and relative experts. Vygotsky (1978) proposed that providing novices with the opportunity to observe experts making use of tools enables novices to develop their knowledge schemas. The AWAR AAR tool suite was observed to encourage senior group members to actively interact with the replay

tools. In line with Vygotsky's idea, junior group members were also observed to take a more active role in elaborating their own understanding of the ABM knowledge domain. A more formulaic and controlled AAR process may not provide as great an opportunity for social learning of this kind.

While these outcomes were positive, several areas for improvement of the EBS10 version of the AWAR AAR tool suite were identified. The first issue is that there was a difference in perspectives between DSTO researchers and the assessor regarding how the AWAR goal hierarchy could be used to inform the AAR discussion. The DSTO scientists expected that the hierarchy scores would be used as a proforma to stimulate discussion points and observations using ground truth replay. Instead, the bookmarked video clips of the mission were used as discussion points, and the hierarchy scores were used as a summary device. A future iteration of the AWAR should examine how accessible the hierarchy of goals and criteria are to the training audience. This may be achieved by refining the structure and content of the hierarchy to more closely match the goals of the training audience.

Another issue with the EBS10 version of the AWAR tool suite was the large amount of feedback it was able to generate. While an exhaustive comparison of objective ground truth with the learners' perceived truth may be of value to the learner, there is a risk of the learners being swamped with feedback. Future iterations of the AWAR tool suite should attempt to summarise the important "take-home" messages of the training mission through the development of new feedback formats.

Lastly, the use of the AWAR tool suite proved relatively labour intensive. At EBS10 one DSTO staff member was devoted to running the tool suite on behalf of the assessor. There may be advantages in modifying the tool suite so that the AAR facilitator may more easily operate it. Some thought should also be given to allowing senior members of the training audience to have some control over the tools to better support audience input into the AAR.

Given that the AWAR was observed to stimulate a more inclusive discussion, what remains to be examined is the processes by which this greater inclusiveness results in training audience learning. How does the assessor providing the AAR actually employ the AWAR tool suite to impact the knowledge of the individual participants in the audience? In Study 2, the Algo-Heuristic Theory (AHT) of learning and instruction (Landa, 1974) is employed to illustrate the differences between a traditional-style squadron debrief and an AAR using the EBS12 version of AWAR. Due to unexpectedly low usage of the AWAR tool suite in EBS12, the data set was constrained to the comparison of a single debrief point using a traditional-style squadron debrief toolset and processes, to a single debrief point employing some of the AWAR capabilities.

3. Study 2

3.1 Algo-Heuristic Theory of learning and instruction

Theories and models of learning and instruction provide insights into how to impart particular types of learning, through appropriate media, using appropriate methods and processes (Gagné, Briggs & Wagner, 1992). The Algo-Heuristic Theory (AHT) of learning and instruction is concerned with identifying and teaching the mental processes that underlie expert performance in any area (Landa, 1983). The AHT approach to learning and instruction has reportedly demonstrated impressive efficiency in the development of expertise in classroom and industrial training settings (Landa, 1994). AHT was developed from AT principles (Bedny & Meister, 1997), and therefore has a theoretical congruence with the analytic lens applied in Study 1. The theoretical congruence is particularly salient with regards to the processes of externalisation/internalisation, as described in AT (Section 2.3).

The processes of externalisation/internalisation are important processes that AAR tool suites such as AWAR may support. For example, an audio-visual representation of a learning event may be used to repair or scale a learner's historical understanding of the event. The audio-visual tool provides support to the cognitive process involved in memory and image or concept manipulation. The audio-visual representation can then be used to modify the learner's understanding of the event to help internalise the new understanding. In developing a better understanding of how the mechanism of externalisation/internalisation works via models such as AHT, better toolsets are able to be designed to support domain specific learning processes.

The AHT approach is a system of techniques for: (1) working with the unobservable, often unconscious mental processes that underlie expert performance, (2) reducing these processes into elementary component operations, (3) developing an algorithmic or heuristic prescription to help the non-expert do in his mind what the expert does, (4) creating a set of learning experiences that allow the non-expert to apply the algorithmic or heuristic prescriptions and develop the automaticity of an expert.

Landa (1975) argues that very often a student grasps the declarative knowledge about a certain subject, but they are not able to apply that knowledge to help them solve problems. For instance, a student may know that in order to swim they need to perform certain movements with their arms and legs, but this does not mean that they are able to swim. They may not understand how to apply these movements to solve the problem of effectively pushing against the water. Similarly, a student may understand the geometric principles of isosceles triangles, but be unable to apply this knowledge in solving related problems of geometry (Landa, 1983). The AHT approach to learning and instruction aims to develop knowledge into knowledge of processes and operations so that they can be applied as skills and abilities. AHT does this by developing the knowledge of processes and operations in the context of problems, which may be made incrementally more general so that the student's understanding becomes more generalised across presentations and contexts.

ABM and Wedgetail controller activity may be broadly construed as a set of spatio-temporal and communication problems. Some of these problems are algorithmic, in that a particular set sequence of steps needs to be undertaken in order to achieve a certain known outcome. Most of the problems are likely to be more heuristic in nature. That is, the controller is working with a more probabilistic set of conditions in a task, where the outcome is not absolutely certain. Landa (1983) asserts that as part of many heuristic tasks there may be some subsets of operations that may be undertaken in an algorithmic way, and this ability is in part what characterises expert performance.

In the context of ABM and Wedgetail Black Skies AARs, there have been examples of persistent problems of performance that have lasted for most of the duration of the exercise. These persistent performance issues have an opportunity-cost associated with them, because they take time to overcome. This lost time could be better used in developing even higher levels of expertise in the training audience. In AHT terms, it is likely that the knowledge that the assessor tries to impart may be understood declaratively by the training audience, but the training audience members are not able to apply this knowledge as a skill or ability. The slow progress in learning outcomes suggests that the knowledge that the assessor tries to impart to the training audience members is not presented in such a way that it is within the ability of the training audience member to accommodate it with their prior understandings. To overcome this issue AHT recommends breaking the knowledge up into mental operations that are sufficiently elementary enough for the training audience member to be able to successfully plan and execute them (externalisation) – and then begin to internalize the method through example problems. This process may require a test of the training audience members' initial level of understanding.

Landa recommends the following 3 general strategies for employing AHT to the instructional/training context (adapted from Landa, 1983): (1) guided discovery, (2) expository teaching and, (3) the combination approach. Guided discovery (strategy 1) gives an overview of the processes involved in the general AHT approach to instruction, including the processes of externalisation/internalisation. The expository teaching and combination approach (strategies 2 and 3) are more practically applicable to the AAR context.

Strategy 1: Guided Discovery
Step 1: Guide the students to discover a system of mental operations underlying a general method of thinking.
Pedagogical actions:
  - *Give the trainees a task or problem and ask them to perform it.*

Step 2: Help the trainee to become aware of what they did in their minds when performing the task and then to formulate a method that corresponds to it.
Pedagogical actions:
  - *Ask them to formulate a set of instructions (a method) so that other people will be able to follow them to perform the task.*
  - *If they have difficulty, explain how to formulate the method.*

- *Identify an overt or hidden (implicit) logical structure of the task or problem content and explicitly describe it.*
- *Show how the logical structure of the content determines the method of handling it.*
- *Use a flowchart when it will help the students to graphically represent the method.*

Step 3: Help the trainee to learn to apply the discovered method.
Pedagogical actions:
- *Have the trainees practice using the discovered method (instructions), in a step-by-step manner, on new cases.*

Step 4: Help the trainee to internalize the method.
Pedagogical actions:
- *Have the trainees practice the method on new cases without looking at instructions and using only self-instructions.*

Step 5: Help the trainee to automatize the method.
Pedagogical actions:
- *Require the trainees to perform the task on new cases very quickly, without using the self-instructions.*

Step 6: Repeat steps 1-5 to increase the degree of generality of the method the trainees have discovered.
Pedagogical actions:
- *In step 1, give the trainees tasks (or problems) that are just outside the subject-matter domain where the method was initially discovered and used, and that require a modification of the discovered method.*
- *In step 2, help the trainees formulate a single more general method that works in both domains.*
- *Steps 3-5 are unchanged.*

Strategy 2 Expository Teaching
The same six steps are undertaken, but the first two are provided to the students in ready-made form.

Strategy 3: Combination Approach
Some steps are taught through discovery and some through expository methods, depending on the instructor/assessor's objectives.

The initial three steps of the AHT approach involve helping the learner externalise their understanding of how to solve a problem, and develop a new or modified approach where required. Step three begins the process of internalisation, moving towards greater automaticity and generality of the knowledge over the remaining steps.

AHT instructional strategies describe a general process of how to use externalisation and internalisation processes to develop training audience member's understanding of how to apply knowledge to tasks and problems. These general processes that AHT employs may also be used as a template to analyse examples of AWAR and non-AWAR-based AAR feedback. This analysis should indicate how AWAR supports an assessor in their approach

to AAR feedback, and how AHT principles may improve the approach to imparting that knowledge. For instance, AWAR hierarchy and ratings may be used to cue the training audience to issues of performance, alternate strategies for which may be developed by expository or guided discovery-type methods (externalisation). This externalised learning may then be applied through some active tool engagement, such as training audience collaboratively developing a diagram on the Smartboard to begin the process of internalisation. Study 1 clearly demonstrated a significant shift in the AAR process with the use of AWAR tools. Would such a shift also be reflected in changes to the pedagogical processes employed by the assessors/debriefers, particularly in terms of externalisation/internalisation processes?

3.2 Method

3.2.1 Participants

The participants involved in this study comprised: (1) a RAAF ABM team consisting of two controllers, a Tactical Director (TD) and an Air Battle Director (ABD) as well as two Air Surveillance Operators; (2) a RAAF Wedgetail mission crew, consisting of two Surveillance and Control Operators (SCOs), a Senior SCO, and a Mission Commander (MC); (3) Seven RAAF simulation pilot personnel (SIMOPs) from Surveillance and Control Training Unit, and three from the School of Air Warfare; (4) a RAAF 41WG (ABM) Assessor, (5) a RAAF 42WG (Wedgetail) Assessor, (6) a RAAF 41WG ABM posted to DSTO as a RAAF Liaison Officer; (7) four ex-military, white-force subject matter experts (SMEs), two of whom were SMEs in ABM operations and the other two were SMEs in F/A-18 Hornet operations.

The ABM and Wedgetail teams were the main training audience, preparing for Exercise Pitch Black 12. The SIMOPs supported the ABM and Wedgetail teams' simulation training by providing simulated fast-jet screen entity and voice interaction. The white-force subject matter experts, consisting of two former ABMs and two former F/A-18 pilots, helped develop and co-ordinate the training scenarios for the training audience. The F/A-18 pilot SMEs worked directly with the SIMOPs during simulation missions to ensure a realistic presentation of entities for the training audience. One of the former-ABM SMEs was responsible for leading the AARs, introducing the participants and maintaining the order of events (See Section 3.2.3).

Assessments of team performance were made in each mission by the 41WG and 42WG Assessors. The 41WG assessor was accompanied in assessment duties by the 41WG RAAF Liaison Officer, or by one of the ABM SMEs. In depth AAR analysis of mission events was undertaken via AWAR by the 41WG and 42WG assessors. The assessors were given a general 'hands-on' introduction to the AWAR software and how it functioned, rather than more formalised training.

UNCLASSIFIED

3.2.2 Materials

3.2.2.1 Video Recordings & AAR Room Layout

One non-AWAR debrief, and 4 AAR sessions using the AWAR review tools were recorded. A representation of the AAR room layout and equipment placement is shown in Figure 6. Three cameras were positioned around the AAR room. One camera was focussed from a gantry over the seated training audience on the AWAR review screen. This camera also recorded audio data from a directional microphone attached separately to the gantry to gather assessor speech from the AWAR review screen area. A second camera was placed at the back of the AAR room to take in the assessor, the AAR display screens, a whiteboard and a rear-view of the training audience. This second camera recorded audio data via a non-directional microphone hung over the training audience from the gantry. A third camera was placed at the front of the AAR room to one side in order to capture the responses of the training audience. This camera also recorded audio data from two directional microphones placed at each front corner of the AAR room to capture training audience verbal responses.
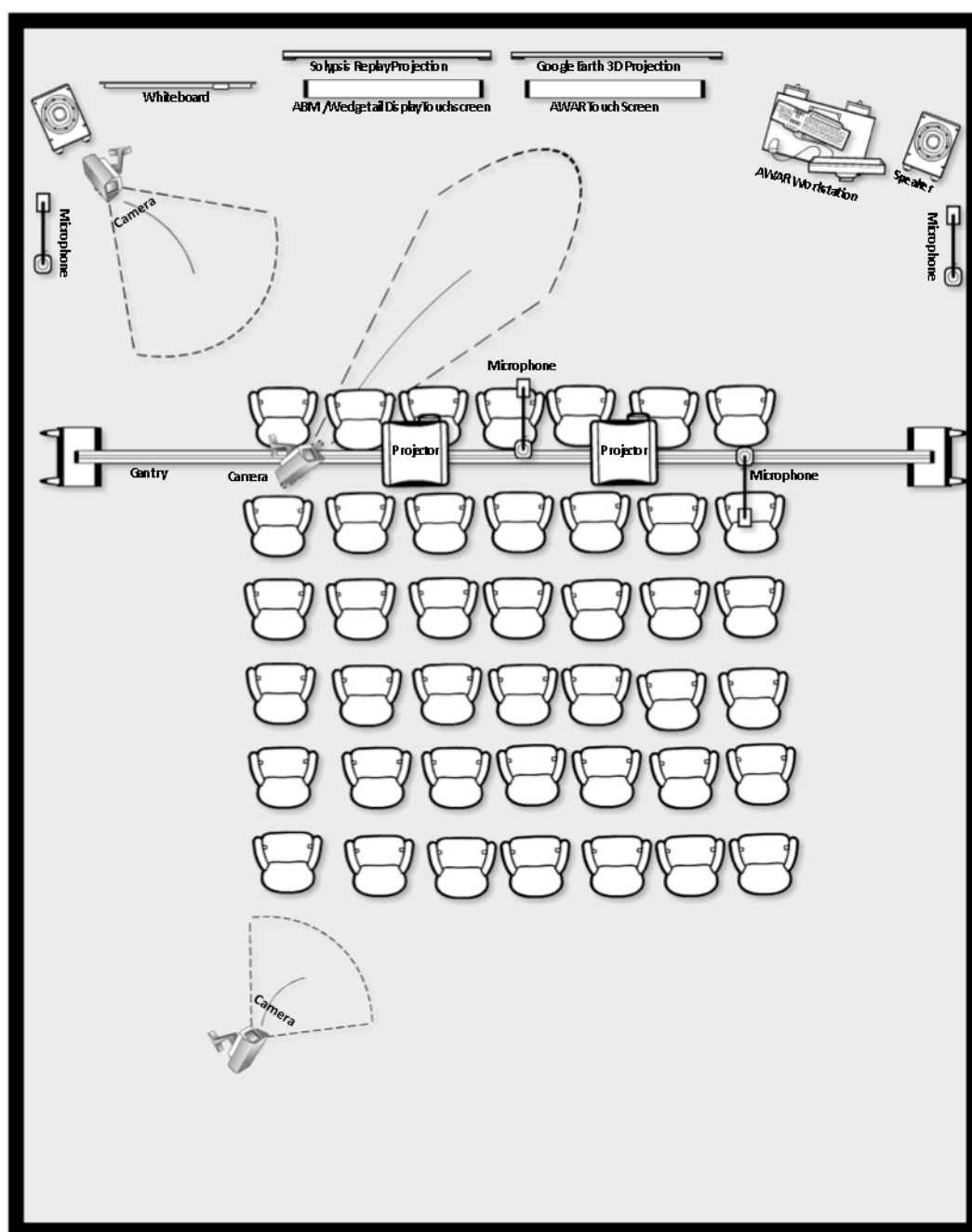
Figure 6. AAR room layout and equipment placement

3.2.2.2 EBS12 Air Warfare Assessment & Review (AWAR) tool suite

The EBS12 version of the AWAR tool suite was similar in terms of its general functionality to the EBS10 version. The major difference in the EBS12 version of AWAR was its development from a spreadsheet-based tool into a more integrated software-based toolset. The differences between the EBS10 and EBS12 versions of AWAR were more marked for the review functionality than the assessment functionality.
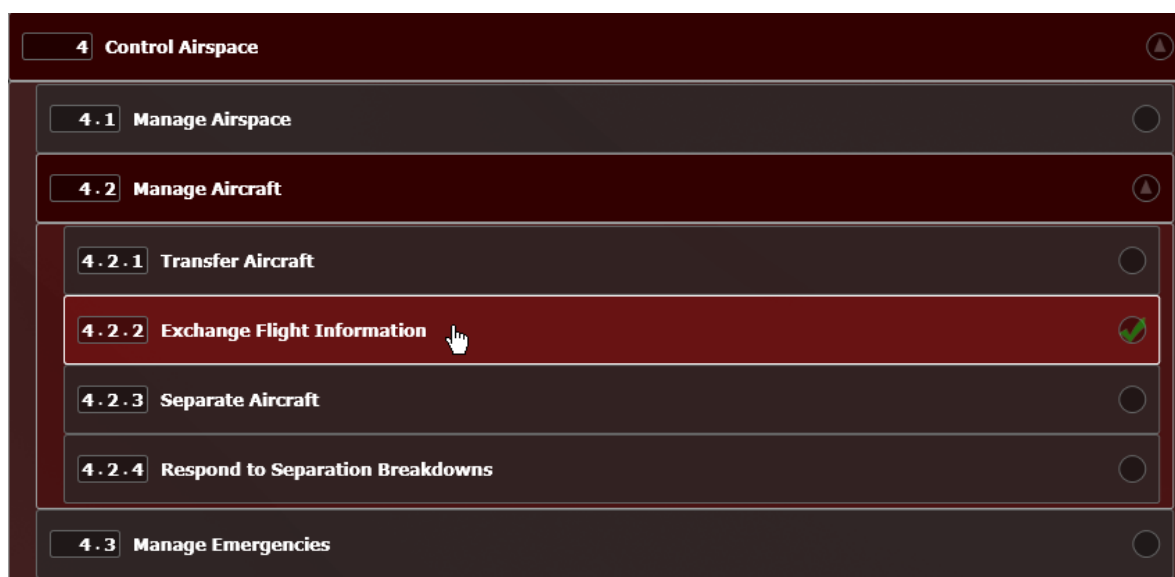
Figure 7 Selection of AWAR assessment hierarchy criteria from EBS12.

The EBS12 version of AWAR was populated with an updated version of the assessment hierarchy. RAAF ABM and Wedgetail assessors made use of the EBS12 AWAR assessment functionality by observing trainees in the simulated mission, and recorded ratings and comments for the team they assessed (Figure 8). As in EBS10, AWAR time-stamped each rating and comment like an internet browser bookmark to help assessors reference those events for audio-visual replay and discussion during the AAR.

During the EBS12 AARs, AWAR generated a review and playback-control interface that the assessor could use to drive the audio-visual replay and encourage discussion (Figure 9). AWAR summarised the ratings given during the assessment phase via a histogram view. To focus on a particular comment or rating, the assessor selected the relevant hierarchy title under the appropriate rating column, which opened a time-stamped point on the video. Bookmark selection and replay was done via the review and playback interface, which aimed to reduce the requirement for assistance from a person managing the AWAR desktop computer to manage selections and screen views.

Figure 8. Review and playback-control interface for AWAR review mode.

The assessors made use of the AWAR review toolset via two Smartboard PG-363 touchscreens. As can be seen in Figure 8, the screen on the right of the AAR room controlled the AWAR review software. The assessor was able to touch the screen in order to select AWAR review content and functionality. The screen on the left of the room presented video recordings of training audience 'scope' content. The scope content was a recording of each ABM or Wedgetail tactical display. The assessor was able to select and expand an individual training audience member's scope replay in order to present feedback. Audio of the relevant comms channels was also available for replay. The scope and audio content was available to examine performance from the perspective of what the ABM or Wedgetail team member saw and heard. Two ground truth displays were also available. Above the individual ABM or Wedgetail tactical displays was the Solipsys display. The Solipsys view was independent of any constraints of the ABM and Wedgetail tactical displays, and showed what the white-force, including the assessors, could see. Over the AWAR review screen was a 3D representation of the ground truth rendered via a modified version of GoogleEarth. This allowed the training audience to see entities represented from a horizontal aerial view in order to be able to observe aircraft in altitude-separation, rather than the top-down view provided by Solipsys. These two ground truth displays were projected onto the wall above the two touchscreens using auditorium projectors.

3.2.3 Procedure

Over the course of EBS12, 5 AARs took place. For each of the AARs, participants gathered in the briefing/debriefing space within several minutes after the afternoon simulated mission each day. The four morning missions involved an abbreviated wrap-up that did not involve AWAR, and were not formally analysed. The initial EBS12 afternoon mission was debriefed in a manner broadly consistent with current squadron practice and was

used as an example of a standard AAR. The agenda of this standard AAR consisted of: (1) an initial introduction by one of the ABM SMEs, (2) an evaluation of the safety and aviation domestics issues that presented in the mission, which was facilitated by one of the F/A-18 SMEs, (3) An examination of a flow (segment) of the mission, again facilitated by the F/A-18 SME, (4) An evaluation of the mission objectives from a 41WG team and 42WG team perspective from the TD and the Senior SCO respectively, (5) An overview of how the AARs would be conducted for the rest of the mission by the ABM SME. No 41WG or 42WG assessor debrief was presented.

Due to the range of expertise involved in EBS12, the core learning event of initial AAR was run by one of the F/A-18 SMEs and focussed largely on tactical execution of the simulated air engagement from the perspective of the fighter aircraft. This AAR focussed on a commentary on one discrete engagement or 'flow' of the mission. The F/A-18 SME used the Solipsys replay as an aid to identifying events and features of the event to provide feedback about.

The remaining AARs made the AWAR tools available to the assessors and followed a consistent agenda. The agenda for these "AWAR AARs" was: (1) an initial introduction by the 42WG MC, (2) a review of 41WG team and 42WG team objects delivered by the TD and the Senior SCO respectively, (3) an evaluation of the safety and aviation domestics issues that presented in the mission, facilitated by the 42WG MC, (4) an AWAR-based debrief facilitated first by the 42WG assessor and then by 41WG assessor, including book-marked video replay and performance evaluation scored against the AWAR hierarchy, where it was considered useful, (5) a traditional evaluation of the performance against the stated 41WG and 42WG objectives facilitated by the TD and Senior SCO respectively, (5) an overview commentary and simulation program notifications provided by an ABM SME.

As in EBS10, the timing of the presentation of the standard AARs before the AWAR AARs was driven by a concerns for order effects with the aim to reduce the possibility of the standard AAR being contaminated by any AWAR-related impacts to AAR style procedures.

3.2.4 Data Analysis

Video material was transferred to an Apple Mac Pro (OS 10.6.7) computer and edited using Final Cut Pro 7. Video data from the three cameras in the AAR sessions was edited to create a synchronised "picture-in-picture" presentation of the three active video camera views. Audio-visual recordings of the entire initial AAR, and the Observer/Assessor AWAR sections of the remaining four AARs were transcribed. These transcriptions not only included the verbal actions of the Observer/Assessors, but also their movements and gestures where these involved engaging with displays to help impart information to the training audience, as per the example below.

Assessor: "I will just…sorry for the re-attack [points towards the relevant comment in the AWAR comment box] on something within the ADGE team today."

A single debrief point from the initial AAR conducted by the F/A-18 SME, and another debrief point from the 41WG assessor that employed AWAR were selected for analysis. These events were nominated by an ABM SME and the RAAF Liaison Officer as being representative of the general 'traditional' approach to AAR. The AWAR AAR debrief point was selected by the researcher due to its relatively greater use of the AWAR capabilities for interaction with the training audience. However, it should be noted that the F/A-18 SME and the 41WG and 42WG assessors do have differing expertise, and this may have shaped their respective approaches to the AAR. Having said this, the experience the F/A-18 SME and assessors bring to the use of the AWAR tool arguably mediates its employment. Despite the SME and assessors having different backgrounds, an analysis of what the AWAR tool suite affords a training audience and how a facilitator may employ AWAR to do this is still informative, although it should be noted that the analysis would only reflect what one facilitator (albeit an experienced one) did.

A step-wise summation of each of these single debrief points was developed from the video transcripts. These step-wise summations describe what the assessor and training audience members did during the interaction with each other and the available feedback tools. These descriptions were augmented with an analysis of the externalisation/internalisation processes that were afforded by the tools available to the assessors, based on the example AHT processes suggested by Landa (1983). The affordances and congruencies of the use of AWAR with Landa's AHT approach to developing skills in students were noted and areas of divergence are discussed below as opportunities for the development of the AWAR AAR procedural workflow.

Due to some technical difficulties that related to AWAR time-synchronisation, a truncated AWAR familiarisation process which likely impacted the judgement of both 41WG and 42WG assessors about when and how to employ the tool suite, many EBS12 AAR debrief points were discussed without reference to AWAR. None of the debrief points that did involve AWAR made use of the audio playback capability, and relied only on video playback. This issue is discussed separately in a report that examines the usability of the EBS12 version of AWAR (Vince & Parker, 2013).

3.3 Results & Discussion

3.3.1 EBS12 Non-AWAR AAR

A step-wise summary of the AAR processes undertaken by the F/A-18 SME is presented in Table 9. Table 9 also presents an analysis of these steps in terms of the learning opportunities interaction with the tools and processes afford the training audience. This AAR point was undertaken in a similar manner to a squadron debrief. The non-AWAR debrief point made use of the Solipsys ground truth replay of an air engagement, but without reference to AWAR scores, comments or audio playback. Due to the background of the SME, tactical fighter execution was the focus of the AAR activity.

Table 9. Non-AWAR AAR feedback summary and analysis

Non-AWAR AAR Step-Wise Summary	Analysis
1. The F/A-18 SME (facilitator) asks the replay assistant to play the Solipsys ground truth video at x2 normal playback rate to advance the video to a point of interest.	The play-through nature of the replay allows the emerging context to be evident to the training audience, unlike a bookmarked approach where the training audience are "dropped" into the learning scenario. Exposure to the emerging context may be useful in some training contexts and irrelevant in others, depending on the relevance of historical factors. (Externalisation)
2. The facilitator gives a description of the 'game-plan' of both fighter teams.	The training audience is given an orientation to the goals of the teams and how they intend to achieve them. This gives the training audience cues to how the event should generally proceed, so that divergences can be noted and discussed. (Externalisation)
3. The facilitator asks for a range-measurement between entities in order to make an evaluation of the relationship between them.	In using the tools available in the Solipsys system to make an evaluation, the facilitator is demonstrating something of how they understand the situation. If the training audience members have the necessary understanding of why he might do this, then it will probably be instructive to them, in terms of an AHT example. However, if the training audience is not able to make sense of this action, then the value as a demonstration is lowered. (Externalisation)
4. The facilitator gives his evaluation of the present actions.	In providing an evaluation of the actions represented through Solipsys, the facilitator provides an opportunity for training audience members who followed the meaning of the range-measurement action to test their assumptions about what measurement meant about the quality of the spatio-temporal relationship between entities. (Externalisation)
5. The facilitator asks for another range-measurement between different entities in order to evaluate the use of airspace – gives an evaluation of spatio-temporal aspects of fighter manoeuvre.	In asking for another range-measurement the facilitator indicates that there is another factor to consider in this situation being replayed, which demonstrates to the training audience how the facilitator analyses the emerging situation. (Externalisation)
6. The facilitator asks for another range between entities.	In providing an evaluation of the actions represented through Solipsys, the facilitator provides another opportunity for training audience members who followed the meaning of the range-measurement action to test their assumptions about what measurement meant about the quality of the spatio-temporal relationship between entities. (Externalisation)
7. A SIMOP asks a question about the realistic presentation of targeting posture for the controller training audience.	The SIMOP role is to provide a realistic presentation of a real fast-jet entity for the ABM & Wedgetail controllers. In asking about this aspect of his role, he is seeking a clarification about his understanding of the situation as it has been presented. (Externalisation)
8. The facilitator answers in terms of time (early) and orientation to target.	The facilitator answers in a way that is relative between entities, rather than an absolute measurement of distance. In doing this, the facilitator has had to unpack the heuristic of his demonstration of his concerns in order to place it within the understanding of the SIMOP. (Externalisation)
9. The facilitator delivers a summary of the important "lessons" in terms of positioning in offensive and defensive orientations.	The facilitator summarises the points of concern for the scenario. Rather than highlighting from the beginning the main points of concern, the summary was given in order bring the observations together in a *post-hoc* way. (Externalisation)

The non-AWAR AAR made no use of bookmarked feedback points, using a full-playback method of presentation. Although some of the playback was truncated through the fast-forward playback function, the process did give the training audience members the possible advantage of a review of historical events. The F/A-18 pilot SME did not undertake any testing of training audience understanding as he undertook his review. It may be a reasonable assumption that the actions he took and feedback points he made were within the range of convenience of the training audience, but this remains unknown. Summary feedback was provided on the basis of *post-hoc* desired outcomes. It was presumed that the training audience would know how these desired outcomes would be best achieved. The process employed by the F/A-18 pilot SME only engaged the process of externalisation of knowledge. There was no strategy for actively applying any of the

feedback through practice examples. The process of internalisation of learning was left to the next mission.

A step-wise summary of the AAR processes undertaken by the 41WG assessor using AWAR is presented in Table 10. Table 10 also presents an analysis of these steps in terms of the learning opportunities interaction with the AAR tools and processes afford the training audience. The AWAR debrief point made use of the Solipsys ground truth replay, rather than the 41WG tactical display of an air engagement, but without reference to AWAR scores, comments or audio playback. Due to the background of the 41WG assessor, ABM skill execution was the focus of the AAR activity.

Table 10. AWAR AAR feedback summary and analysis

AWAR AAR Step-Wise Summary	Analysis
1. The 41WG assessor drags the Solipsys ground truth display down to the Smartboard and places it over the ABM team scopes.	The 41WG assessor wants to access the ground truth display instead of the individual ABM scope views. This may be a cue to the type of information about to be presented.
2. The 41WG assessor describes the air engagement scenario in relation to the history of performance through EBS12 to that point, both in terms of what has been working well and what hasn't worked well.	The assessor is making explicit the areas of performance that have been adequate, and cueing the training audience to a persistent problem that he wants to attend to in order to fix. (Externalisation)
3. The 41WG assessor names the main issue he wants to focus on, and opens the relevant AWAR point. Solipsys opens at a useful point – the 41WG assessor outlines the specific scenario with reference to the entities and their spatio-temporal positions on the display.	The AWAR review screen label indicates the type of issue it is. The rating attached to the selected event indicates how it was viewed by the 41WG assessor. These aspects are cues for the training audience which they may use to match to their memory of events. It is unknown to what degree the training audience noted the AWAR label, as the assessor referred to the event in the scenario context, rather than as a hierarchy label. Rarely in EBS12 did either of the 41WG and 42WG assessors refer to the rating they gave an assessed event. The 41WG assessor then uses the Solipsys display to define the scenario presented on the screen in terms of a category of performance. The training audience members are presented with an opportunity to check their categorisation of the scenario event against this label and description. (Externalisation)
4. The 41WG assessor asks a particular ABM controller about their understanding of a principle of the ABM role.	In asking a training audience member about their understanding of a principle, the assessor is able to check the understanding of that training audience member with reference to the scenario as it was described. The assessor then has the opportunity to correct or generalise the understanding for application to the scenario under review and others that may be considered in the same category. This process allows other audience members to check their own understanding and assess, (1) whether the concept is one that they thought to be important to raise in the scenario context, and (2) whether their understanding aligns with that of the ABM controller. (Externalisation)
5. The training audience member answers the assessor firstly in terms of the general principle, and then gives a description of how they think about that general principle in the given scenario.	The training audience member indicates that he understands that the application of the general principle may differ depending on circumstances. (Externalisation)
6. The 41WG assessor asks a Wedgetail controller about the basic role of an asset in the air engagement.	In asking a training audience member about their understanding of an asset's role, the assessor is able to check the understanding of that training audience member with reference to the scenario as it was described. The assessor then has the opportunity to correct or generalise the understanding for application to the scenario under review and others that may be considered in the same category. This process allows other audience members to check their own understanding and assess, (1) whether the concept is one that they thought to be important to raise in the scenario context, and (2) whether their understanding aligns with that of the ABM controller. (Externalisation)
7. The Wedgetail controller answers the question about the basic role of the particular air asset, without making an attempt to relate that principle to the scenario under discussion.	The controller is able to answer the question about the basic role, but appears unable to communicate an understanding of what this asset's basic role may mean in relation to the scenario and the principle previously discussed by the ABM controller. (Externalisation)
8. The 41WG assessor links the role of the asset to a better known more basic concept.	The assessor has drawn the parallel between the advanced knowledge he is trying to convey and a set of operations that are already very well known to the training audience. (Externalisation)
9. The 41 WG assessor demonstrates how the concept works on the Solipsys screen by using the Smartboard pens to draw the impact of the concept over the current scenario.	The assessor externalises the links between the declared principle, as it relates to the asset, using as an analogy a process the training audience already know well. The Smartboard allows these concepts to be externally represented. (Externalisation)
10. The 41WG assessor points back to the historical problems that could have been solved by the adoption of this approach.	11. Ideally, AWAR would allow these instances to be recalled, or allow them to be replayed for purposes of generalisation. Generalisation necessarily follows in future missions, if the feedback has been effective.

The AWAR AAR made use of the bookmark function of AWAR, which enabled the training audience to be 'dropped in' to a scenario for evaluation. This approach saves time over replaying unnecessary video, but relies on the assessor to make an informative

introduction to the scenario under consideration. AWAR provided cues to the training audience about the nature of the issues under review, according to the assessment hierarchy and the attached rating. This presented an opportunity for the training audience members to check whether their experience of the mission event corresponded with that of the assessor.

The assessor was able to use Solipsys to give a description of the salient elements in the picture and their relevant qualities. The assessor used questions to check the level of understanding of some of the training audience members. This is an important step because, as AHT suggests, if the mental operations to be learned are not simple enough, future practice will be based on a process more akin to trial and error. The assessor invoked processes that were well known to the training audience members in order to make what he was trying to convey within the range of convenience of the training audience members' experience. The assessor's use of the Smartboard to illustrate the learning point was useful in creating a physical representation to support the verbal representation for the training audience. This continued the process of externalisation. According to AHT, if the assessor asked a training audience member to use the Smartboard to apply the knowledge he had described, this would have advanced the process of internalisation for the next mission.

The analysis of the non-AWAR and AWAR AAR points indicates that there were no opportunities to undertake internalisation processes (practice of feedback concepts) prior to the next training mission. Indeed, the AHT processes of internalisation, automatisation and generalisation of training audience mental operations were left to *ad-hoc* mission practice. There were differences in information available to the training audience, such as the hierarchy label, attached comment and rating. However, the 41WG assessor did not formally point out these information sources. Audio playback was also available, but was either judged irrelevant to the learning point, or was deemed as too difficult to incorporate into the debrief point discussion. It appears that despite the capabilities provided by the AWAR toolset, the assessor conducted an AAR that mirrored the main steps of a standard AAR, not moving beyond the externalisation steps. The assessor had had limited training with the toolset so may not have been confident or comfortable to explore the full set of options available – preferring to only augment their standard AAR process 'around the edges'. The provision of advanced toolsets does not guarantee they will be used optimally. This speaks to the need to adequately 'train the trainer'.

There remains a question as to whether there is a value in extending the AAR toolset so that it affords some modes of practice during the AAR itself. These methods may include, for example: (1) event walk-throughs of similar scenarios using historical or tool-contrived media, (2) Smartboard mud-maps and illustrations, (3) manipulating bats (aircraft models), and (4) comms role-plays. While these activities may not be a part of the present AAR agenda as a result of time constraints, time may arguably be saved if the training audience is able to enter the next training mission having begun the internalisation process.

Tools such as AWAR will arguably not make a poorly skilled assessor as effective as a skilled one. AWAR is still reliant on the skills, experience and insight of an assessor. In

particular, skilled assessors engage in processes such as checking the understanding of training audience members and packaging information so that it can be internalised and applied.

4. Conclusion

The present analysis of the impact on training of the EBS10 and EBS12 versions of the AWAR AAR tool suite, was made up of two related studies. Study 1 focused on the pattern of interaction between the assessor leading the AAR and the training audience as mediated by the EBS10 AWAR toolset. EBS10 AAR data was analysed through the lens of Activity Theory (AT), which is useful in examining how tools mediate social and cognitive processes (Leont'ev, 1978). Study 1 found that: (1) the use of the AWAR enabled a more inclusive and discursive AAR than the non-AWAR AAR, (2) AWAR provided a useful objective ground truth to participants, and (3) AWAR opened opportunities for social learning via the observation of experts.

Key issues that arose from the introduction of AWAR included: (1) AWAR generated large volumes of feedback for the training audience to absorb, (2) AWAR was labour intensive to run as a spread-sheet database tool.

Study 1 focussed on the patterns of interaction in AWAR and non-AWAR AARs, and did not examine assessor processes of feedback in a very detailed manner.

The EBS12 version of AWAR was developed as a software-based tool, which addressed the issues of the EBS10 version by: (1) highlighting the use of bookmarks to allow the assessors to select particular sections of a mission for analysis, rather than a whole mission; (2) giving the assessor bookmark selection and playback controls, so that they could run the AAR with a reduced level of assistance. The usability of these changes are examined in Vince & Parker (2013).

Study 2 examined the AAR feedback process at the level of discussion of a single learning point. Study 2 focused on the strategies used by the assessor to develop the knowledge of the training audience as mediated by the AWAR toolset. EBS12 AAR data was analysed with reference to the Algo-Heuristic Theory (AHT) of learning and instruction (Landa, 1974). AHT is a theory of learning and instruction aimed at the development of both expertise and general cognitive skills (Bedny & Meister, 1997).

Key issues that arose from the analysis undertaken in Study 2 are: (1) while the AWAR allows for an objective ground truth to be represented, the feedback that is developed from this ground truth is reliant on the skills of the assessor, particularly in terms of checking the understanding of the training audience, and (2) the feedback and learning process may arguably benefit from the inclusion of some practice exercises prior to the following mission.

There are two recommendations that arise from these issues: (1) it may be of benefit to trial the effectiveness of the AHT approach to developing expertise in a laboratory setting, and (2) if these trials of AHT processes are effective, AWAR or EBS14 AAR procedures may be modified to include some practice exercises in line with AHT recommendations.

5. Acknowledgements

6. References

Annett, J., Cunningham, D., & Mathias-Jones, P. (2000). A method for measuring team skills. *Ergonomics*, 43 (8), 1076-1094.

Barab, S., A., Barnett, M., Yamagata-Lynch, L., Squire, K. & Keating, T. (2002). Using activity theory to understand the contradictions characterizing a technology-rich introductory astronomy course. *Mind, Culture, and Activity*, 9(2), 76-107.

Basharina, O. K. (2007). An activity theory perspective on student-reported contradictions in international telecollaboration. *Language Learning & Technology*, 11(2), 82-103. [viewed 6th October 2011]. http://llt.msu.edu/vol11num2/basharina/default.html

Bedny, G. Z., & Meister, D. (1997). *The Russian Theory of Activity: Current Applications to Design and Learning*. Mahwah, NJ: Lawrence Erlbaum Associates.

Bellamy, R. (1996). Designing educational technology: Computer-mediated change. In B. Nardi (Ed.), *Context and consciousness: Activity theory and human-computer interaction* (pp. 123-145). Cambridge: The MIT Press.

Best, C., Hasenbosch, S., Skinner, M., Crane, P., Burchat, E., Finch, M., Gehr, S., Kam, C., Shanahan, C., & Zamba, M. (2007). Exercise Pacific Link 2: Coalition Distributed Mission Training Research. In *Proceedings of the SimTecT 2007 Conference.* Brisbane, Australia: Simulation Industry Association of Australia.

Engeström, Y. (1987). *Learning by expanding: An activity-theoretical approach to developmental research*. Helsinki: Orienta-Konsultit.

Gagné, R.M. , Briggs, L.J. & Wager, W.W. (1992). *Principles of instructional design.* New York: Holt, Rinehart & Winston.

Glaser, B. & Strauss, A. (1967). The discovery of grounded theory. Chicago: Aldine.

Goldberg, S.L. and Meliza, L.L. (1993). Assessing unit performance in distributive interactive simulations: The Unit Performance Assessment System (UPAS). *Proceedings of NATO Defence Research Group Meeting, Panel 8 (Defence Applications of Human and Bio-Medical Sciences). Training Strategies for Networked Simulation and Gaming). Technical Proceedings AC/243 (Panel 8) TN/5*, 173-182.

Hasenbosch, S. & Best, C. (2007). *A Hierarchical Analysis of Air Battle Management Team Goals in the Defensive Counter Air Mission.* DSTO-TN-0781. Melbourne: Defence Science and Technology Organisation.

Hoare, R. (1996). From debrief to After Action Review (AAR*). Modern Simulation & Training*, 6, 13-17.

Kaptelinin, V., Nardi, B. & Macaulay, C. (1999). The activity checklist: A tool for representing the "space" of context. *Interactions*, 6(4), 27-39.

Landa, L.N. (1974). *Algorithmization in learning and instruction.* Englewood Cliffs: Educational Technology Publications.

Landa. L.N. (1975). Some problems in algorithmization and heuristics in instruction. *Instructional Science, 4*, 99-112.

Landa, L.N. (1983). The algo-Heuristic Theory of Instruction. In C.M. Reigluth, (Ed.), *Instructional design theories and models: An overview of their current status.* Hillsdale: Lawrence Earlbaum Associates.

Landa, L.N. (1994). The creation of expert performers without years of conventional experience: The landamatics method. *The Journal of Management Development, 6* (4), 40-52.

Leont'ev, A.N. (1978). *Activity, Consciousness, Personality.* Englewood Cliffs: Prentice Hall.

McCarthy, A., Kingston, G., Johns, K., Gori, R., Main, P. & Kruzins, E. (2003). *Joint Warfare Capability Assessment - Final Report: Australian Joint Essential Tasks Volume 1.* DSTO-CR-0293. Melbourne: Defence Science and Technology Organisation.

Meliza, L.L. (1999). *A guide to standardizing after action review (AAR) aids* (ARI Research Product 99-01). Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences.

Meliza, L.L., Golberg, S.L., & Lampton, D.R. (2008). *After action review in simulation-based training.* RTO-TR-HFM-121-Part II. North Atlantic Treaty Organisation (NATO), Research and Technology Organisation (RTO): Paris, France

Murphy, E. & Rodriguez-Manzanares, M.A. (2008). Using activity theory and its principle of contradictions to guide research in educational technology. *Australasian Journal of Educational Technology*, 24(4), 442-457.

Paraskevas, A. & Wickens, E. (2003) Andragogy and the Socratic method: The adult learner persective. *Journal of Hospitality, Leisure, Sport & Tourism Education,* 2(2), 4-14.

Shanahan, C., Best, C., Finch, M., Stott, A., Tracey, E., & Hasenboch, S (2009). Exercise Black Skies 2008: Enhancing Live Training Through Virtual Preparation Part One: An Evaluation of Training Effectiveness. DSTO-RR-0344. Melbourne: Defence Science and Technology Organisation.

Shute, V.J. (2008). Focus on formative feedback. *Review of Educational Research,* 78(1), 153-198.

Stevens, A., Crone, D., Temby, P., Best, C., & Simpkin, G. (2011, May). *Using synthetic environments to enhance close air support training.* Paper presented at the Simulation Conference & Exhibition (SimTecT - 2011), Melbourne.

Stott, A., Best, C., & Shanahan, C. (2011, May). *Enhancing standardisation of Air Battle Manager team performance assessment.* Paper presented at the Simulation Conference & Exhibition (SimTecT - 2011), Melbourne.

Strong, M. (1996) *The Habit of Thought: From Socratic Seminars to Socratic Practice*. Chapel Hill, NC: New View Publications.

Tredway, L. (1995). Socratic Seminars: Engaging Students in Intellectual Discourse. *Educational Leadership,* 53(1), 26-29.

Vygotsky, L.S. (1978). *Mind in Society*. Cambridge: Harvard University Press.

DEFENCE SCIENCE AND TECHNOLOGY ORGANISATION DOCUMENT CONTROL DATA		1. DLM/CAVEAT (OF DOCUMENT)	
2. TITLE An Evaluation of New After-Action Review Tools in Exercise Black Skies 10 & Exercise Black Skies 12		3. SECURITY CLASSIFICATION (FOR UNCLASSIFIED REPORTS THAT ARE LIMITED RELEASE USE (L) NEXT TO DOCUMENT CLASSIFICATION) Document (U) Title (U) Abstract (U)	
4. AUTHOR(S) Julian Vince and Christopher Best		5. CORPORATE AUTHOR DSTO Defence Science and Technology Organisation 506 Lorimer St Fishermans Bend Victoria 3207 Australia	

6a. DSTO NUMBER DSTO-TN-1216	6b. AR NUMBER AR-015-746	6c. TYPE OF REPORT Technical Note	7. DOCUMENT DATE October 2013

8. FILE NUMBER 2011/1216837/1	9. TASK NUMBER AIR07/327	10. TASK SPONSOR DGSP-AF	11. NO. OF PAGES 36	12. NO. OF REFERENCES 27

13. DSTO Publications Repository http://dspace.dsto.defence.gov.au/dspace/	14. RELEASE AUTHORITY Chief, Aerospace Division

15. SECONDARY RELEASE STATEMENT OF THIS DOCUMENT
Approved for public release OVERSEAS ENQUIRIES OUTSIDE STATED LIMITATIONS SHOULD BE REFERRED THROUGH DOCUMENT EXCHANGE, PO BOX 1500, EDINBURGH, SA 5111

16. DELIBERATE ANNOUNCEMENT
No Limitations

17. CITATION IN OTHER DOCUMENTS	Yes

18. DSTO RESEARCH LIBRARY THESAURUS
Training Research, After-Action Review, Tools, Evaluation, Activity Theory, Algo-Heuristic Theory

19. ABSTRACT
An evaluation of a new suite of after action review tools was undertaken as part of a series of synthetic collective training research exercises (Exercise Black Skies 10 & Exercise Black Skies 12). This technical note provides a description of the observed benefits of this tool suite and proposes opportunities for future development.